



**ANALISIS PERBANDINGAN ALGORITMA *K-MEANS* DAN
HIERARCHICAL CLUSTERING UNTUK SEGMENTASI PELANGGAN
DALAM OPTIMALISASI STRATEGI PEMASARAN MENGGUNAKAN
METODE *RECENCY, FREQUENCY, MONETARY* (STUDI KASUS: PT
XYZ)**

LAPORAN TUGAS AKHIR

HALAMAN J

RIDWAN SAPUTRA SOFYANSAH

41521010078

UNIVERSITAS
MERCU BUANA

**PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS ILMU KOMPUTER
UNIVERSITAS MERCU BUANA
JAKARTA
2025**



**ANALISIS PERBANDINGAN ALGORITMA *K-MEANS* DAN
HIERARCHICAL CLUSTERING UNTUK SEGMENTASI PELANGGAN
DALAM OPTIMALISASI STRATEGI PEMASARAN MENGGUNAKAN
METODE *RECENCY, FREQUENCY, MONETARY* (STUDI KASUS: PT
XYZ)**

LAPORAN TUGAS AKHIR

ALAMAN

**RIDWAN SAPUTRA SOFYANSAH
41521010078**

UNIVERSITAS
MERCU BUANA

Diajukan sebagai salah satu syarat untuk memperoleh gelar sarjana

**PROGRAM STUDI TEKNIK INFORMATIKA
FAKULTAS ILMU KOMPUTER
UNIVERSITAS MERCU BUANA
JAKARTA
2025**

HALAMAN PERNYATAAN KARYA SENDIRI

Saya yang bertanda tangan di bawah ini:

Nama : Ridwan Saputra Sofyansah
NIM : 41521010078
Program Studi : Teknik Informatika
Judul Laporan Skripsi : Analisis Perbandingan Algoritma K-Means dan Hierarchical Clustering untuk Segmentasi Pelanggan dalam Optimalisasi Strategi Pemasaran Menggunakan Metode Recency, Frequency, Monetary (Studi Kasus: PT XYZ)

Menyatakan bahwa Laporan Skripsi ini adalah hasil karya saya sendiri dan bukan plagiat, serta semua sumber baik yang dikutip maupun dirujuk telah saya nyatakan dengan benar. Apabila ternyata ditemukan di dalam Laporan Skripsi saya terdapat unsur plagiat, maka saya siap mendapatkan sanksi akademis yang berlaku di Universitas Mercu Buana.

Jakarta, 19 Juli 2025



Ridwan Saputra Sofyansah

HALAMAN PENGESAHAN

Laporan Skripsi ini diajukan oleh:

Nama : Ridwan Saputra Sofyansah
NIM : 41521010078
Program Studi : Teknik Informatika
Judul Laporan Skripsi : Analisis Perbandingan Algoritma K-Means dan Hierarchical Clustering untuk Segmentasi Pelanggan dalam Optimalisasi Strategi Pemasaran Menggunakan Metode Recency, Frequency, Monetary (Studi Kasus: PT XYZ)

Telah berhasil dipertahankan pada sidang di hadapan Dewan Penguji dan diterima sebagai bagian persyaratan yang diperlukan untuk memperoleh gelar Sarjana Strata 1 pada Program Studi Teknik Informatika, Fakultas Ilmu Komputer Universitas Mercu Buana.

Disahkan oleh:

Pembimbing : Muhammad Rifqi, S.Kom., M.Kom
NIDN : 0301067101
Ketua Penguji : Inna Sabily Karima, S.Kom, M.Kom
NIDN : 0324018902
Penguji 1 : Rushendra, S.Kom, MT
NIDN : 0408067402
Penguji 2 : Prastika Indriyanti, S.Kom, MCS
NIDN : 0312089401

Jakarta, 19 Juli 2025

Mengetahui,

Dekan

Ketua Program Studi

Dr. Bambang Jokonowo, S.Si., MTI
NIDN : 0320037002

Dr. Hadi Santoso, S.Kom., M.Kom
NIDN : 0225067701

KATA PENGANTAR

Puji syukur kehadiran Tuhan yang Maha Esa, atas segala rahmat dan ridha-Nya sehingga penulis dapat menyelesaikan tugas akhir yang merupakan salah satu persyaratan kelulusan Program Studi Strata Satu (S1) pada jurusan Teknik Informatika, Universitas Mercu Buana.

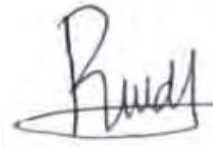
Penulis menyadari bahwa tugas akhir ini masih jauh dari sempurna, karena kesempurnaan sejatinya hanya milik Tuhan yang Maha Esa. Oleh karena itu, saran dan masukan yang membangun senantiasa penulis terima dengan senang hati. Serta berkat dukungan, motivasi, bantuan, bimbingan, dan doa dari banyak pihak, penulis mengucapkan terima kasih kepada:

1. Bapak Prof. Dr. Andi Adriansyah, M.Eng. selaku Rektor Universitas Mercu Buana.
2. Bapak Dr. Bambang Jokonowo, S.Si., MTI selaku Dekan Fakultas Ilmu Komputer.
3. Bapak Dr. Hadi Santoso, S.Kom., M.Kom. selaku Ketua Program Studi Teknik Informatika Universitas Mercubuana.
4. Bapak Muhammad Rifqi, S.Kom., M.Kom. selaku dosen pembimbing Tugas Akhir yang telah memberikan pengarahan serta meluangkan waktu, tenaga, dan pikiran selama proses penyusunan tugas akhir.
5. Seluruh dosen Fakultas Ilmu Komputer Universitas Mercu Buana yang telah memberikan ilmu serta wawasan yang sangat berharga selama masa perkuliahan.
6. PT XYZ yang telah memfasilitasi penulis dengan memberikan izin serta menyediakan data transaksi yang diperlukan sebagai bahan penelitian.
7. Kedua Orang Tua dan kedua kakak kandung penulis yang senantiasa memberikan dukungan selama menjalani masa studi sebagai mahasiswa Universitas Mercu Buana.
8. Sahabat sejak kecil penulis, Sinta Fertika Sari, yang telah memberikan semangat, bantuan, serta menjadi tempat berbagi dalam setiap proses yang dilalui selama penyusunan tugas akhir ini.
9. Rekan-rekan mahasiswa yang telah berbagi informasi dan memberikan dukungan dalam berbagai bentuk.

10. Diri penulis sendiri, atas ketekunan, semangat, dan komitmen yang terus dijaga dalam menghadapi setiap tantangan selama proses perkuliahan hingga penyusunan tugas akhir ini.

Akhir kata, penulis berharap semoga Tuhan yang Maha Esa membalas kebaikan dan selalu mencurahkan rahmat, hidayah, serta panjang umur kepada kita semua, aamiin. Terima Kasih.

Jakarta, 19 Juli 2025



Ridwan Saputra Sofyansah



**HALAMAN PERNYATAAN PERSETUJUAN PUBLIKASI TUGAS
AKHIR UNTUK KEPENTINGAN AKADEMIS**

Sebagai sivitas akademik Universitas Mercu Buana, saya yang bertanda tangan di bawah ini:

Nama : Ridwan Saputra Sofyansah
NIM : 41521010078
Program Studi : Teknik Informatika
Judul Laporan Skripsi : Analisis Perbandingan Algoritma K-Means dan Hierarchical Clustering untuk Segmentasi Pelanggan dalam Optimalisasi Strategi Pemasaran Menggunakan Metode Recency, Frequency, Monetary (Studi Kasus: PT XYZ)

Demi pengembangan ilmu pengetahuan, dengan ini memberikan izin dan menyetujui untuk memberikan kepada Universitas Mercu Buana **Hak Bebas Royalti Non-Eksklusif (*Non-exclusive Royalty-Free Right*)** atas karya ilmiah saya yang berjudul di atas beserta perangkat yang ada (jika diperlukan).

Dengan Hak Bebas Royalti Non-Eksklusif ini Universitas Mercu Buana berhak menyimpan, mengalihmedia/format-kan, mengelola dalam bentuk pangkalan data (*database*), merawat, dan mempublikasikan Laporan Magang/Skripsi/Tesis/Disertasi saya selama tetap mencantumkan nama saya sebagai penulis/pencipta dan sebagai pemilik Hak Cipta.

Demikian pernyataan ini saya buat dengan sebenarnya.

Jakarta, 19 Juli 2025

Yang menyatakan,



Ridwan Saputra Sofyansah

ABSTRAK

Nama : Ridwan Saputra Sofyansah
NIM : 41521010078
Program Studi : Teknik Informatika
Judul Laporan Skripsi : Analisis Perbandingan Algoritma K-Means dan Hierarchical Clustering untuk Segmentasi Pelanggan dalam Optimalisasi Strategi Pemasaran Menggunakan Metode Recency, Frequency, Monetary (Studi Kasus: PT XYZ)
Dosen Pembimbing : Muhammad Rifqi, S.Kom., M.Kom

Di tengah meningkatnya persaingan bisnis khususnya di bidang retail, segmentasi pelanggan menjadi salah satu strategi kunci untuk mengenali kelompok pelanggan berdasarkan pola pembelian mereka. Tujuan dari studi ini yaitu untuk membandingkan efektivitas dan efisiensi algoritma K-Means dan Hierarchical Clustering dalam mengidentifikasi kelompok pelanggan, menggunakan pendekatan Recency, Frequency, dan Monetary yang didasarkan pada data transaksi pelanggan PT XYZ sepanjang tahun 2024. Studi ini menerapkan pendekatan kuantitatif komparatif, mencakup tahapan eksplorasi dan pembersihan data, perhitungan skor RFM, normalisasi data, penerapan algoritma clustering, serta evaluasi kinerja menggunakan metrik Silhouette Score, Davies-Bouldin Index, Cophenetic Correlation Coefficient, dan waktu komputasi. Temuan studi ini mengindikasikan bahwa metode K-Means mampu memberikan hasil nilai Silhouette Score sebesar 0,6505 dan Davies-Bouldin Index sebesar 0,4515 dengan waktu komputasi 0,0452 detik, sementara Hierarchical Clustering menghasilkan nilai Silhouette Score sebesar 0,6354, Davies-Bouldin Index sebesar 0,4647, dan Cophenetic Correlation Coefficient sebesar 0,8469 dengan waktu komputasi 46,8623 detik. Segmentasi pelanggan menghasilkan tiga kelompok utama, yaitu Best Customers (pelanggan aktif dengan frekuensi tinggi dan kontribusi transaksi besar), Promising Customers (pelanggan dengan potensi tinggi namun aktivitas sedang), dan At-Risk Customers (pelanggan yang jarang bertransaksi dan memiliki kontribusi rendah). Dari hasil pengujian, metode K-Means terbukti memberikan hasil yang lebih optimal dan bekerja lebih cepat, sehingga direkomendasikan untuk segmentasi pelanggan berbasis RFM. Hasil segmentasi digunakan untuk menyusun pendekatan pemasaran yang lebih terfokus, seperti penyusunan program loyalitas bagi pelanggan bernilai tinggi serta strategi retensi untuk pelanggan yang berpotensi berhenti atau tidak lagi bertransaksi, sehingga memperkuat proses pengambilan keputusan yang didasarkan pada analisis data di PT XYZ.

Kata kunci: Segmentasi Pelanggan, K-Means, Hierarchical Clustering, RFM, Retail

ABSTRACT

Name : Ridwan Saputra Sofyansah
NIM : 41521010078
Study Program : Teknik Informatika
Title Thesis : Analisis Perbandingan Algoritma K-Means dan Hierarchical Clustering untuk Segmentasi Pelanggan dalam Optimalisasi Strategi Pemasaran Menggunakan Metode Recency, Frequency, Monetary (Studi Kasus: PT XYZ)
Counsellor : Muhammad Rifqi, S.Kom., M.Kom

Amidst the intensifying competition in the retail industry, customer segmentation has become a key strategy to identify customer groups based on their purchasing patterns. This study aims to compare the effectiveness and efficiency of the K-Means and Hierarchical Clustering algorithms in identifying customer segments using the Recency, Frequency, and Monetary approach, based on PT XYZ's customer transaction data throughout 2024. A comparative quantitative approach is applied, including stages of data exploration and cleaning, RFM scoring, data normalization, clustering algorithm implementation, and performance evaluation using metrics such as Silhouette Score, Davies-Bouldin Index, Cophenetic Correlation Coefficient, and computational time. The findings of this study indicate that the K-Means method achieved a Silhouette Score of 0,6505 and a Davies-Bouldin Index of 0,4515 with a computation time of 0,0452 seconds, whereas the Hierarchical Clustering method yielded a Silhouette Score of 0,6354, a Davies-Bouldin Index of 0,4647, and a Cophenetic Correlation Coefficient of 0,8469 with a computation time of 46,8623 seconds. The segmentation process produced three main customer groups: Best Customers (highly active customers with significant transaction contributions), Promising Customers (potentially valuable customers with moderate activity), and At-Risk Customers (infrequent buyers with low contribution). Based on the evaluation, the K-Means method demonstrated more optimal results and faster performance, making it more suitable for RFM-based customer segmentation. The segmentation outcomes are utilized to design more targeted marketing approaches, such as loyalty programs for high-value customers and retention strategies for those at risk of churning, thereby enhancing data-driven decision-making processes at PT XYZ.

Keywords: Customer Segmentation, K-Means, Hierarchical Clustering, RFM, Retail

DAFTAR ISI

HALAMAN JUDUL	i
HALAMAN PERNYATAAN KARYA SENDIRI	ii
HALAMAN PENGESAHAN.....	iii
KATA PENGANTAR.....	iv
HALAMAN PERNYATAAN PERSETUJUAN PUBLIKASI TUGAS AKHIR UNTUK KEPENTINGAN AKADEMIS	vi
ABSTRAK	vii
ABSTRACT.....	viii
DAFTAR ISI.....	ix
DAFTAR TABEL	xii
DAFTAR GAMBAR.....	xiii
DAFTAR LAMPIRAN.....	xiv
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang	1
1.2 Perumusan Masalah	3
1.3 Tujuan Penelitian	4
1.4 Manfaat Penelitian	4
1.5 Batasan Masalah.....	5
BAB II TINJAUAN PUSTAKA.....	6
2.1 Penelitian Terdahulu	6
2.2 Teori Pendukung.....	14
2.2.1 Segmentasi Pelanggan	14
2.2.2 Algoritma Clustering	15

2.2.3 RFM.....	17
2.2.4 Industri Ritel.....	17
2.2.5 Customer Relationship Management (CRM).....	18
2.2.6 Strategi Pemasaran	18
2.2.7 Data Mining.....	19
2.2.8 Pemrograman Komputer.....	19
2.2.9 Analisis Perbandingan (Komparatif)	20
2.3 Perbandingan dengan Baseline Penelitian	21
2.4 Gap Penelitian	23
BAB III METODE PENELITIAN	25
3.1 Jenis Penelitian.....	25
3.2 Tahapan Penelitian	25
BAB IV HASIL DAN PEMBAHASAN	30
4.1 Deskripsi Dataset	30
4.2 Proses Modeling Machine Learning untuk Clustering	32
4.2.1 Data Exploration.....	32
4.2.2 Data Cleaning	34
4.2.3 Data Preprocessing / Feature Engineering.....	35
4.2.4 Modeling.....	39
4.2.5 Evaluasi Clustering.....	50
4.2.6 Analisis Sensitivitas Parameter Clustering.....	52
4.3 Hasil Analisis	53
4.3.1 Hasil Segmentasi dengan K-Means.....	53
4.3.2 Hasil Segmentasi dengan Hierarchical Clustering	55
4.3.3 Hasil Evaluasi Performa Model Clustering	56
4.3.4 Perbandingan Waktu Komputasi.....	56
4.4 Pembahasan.....	57
4.4.1 Perbandingan Hasil Clustering K-Means dan Hierarchical.....	57
4.4.2 Interpretasi Segmentasi Pelanggan.....	58
4.4.3 Implikasi Strategis	60
BAB V KESIMPULAN DAN SARAN	62

5.1	Kesimpulan	62
5.2	Saran.....	63
DAFTAR PUSTAKA.....		66
LAMPIRAN.....		71



DAFTAR TABEL

Table 2.1 Penelitian Terdahulu	6
Table 2.2 Perbandingan Penelitian Ini dengan Baseline Penelitian.....	21
Table 2.3 Perbandingan Penelitian Terdahulu untuk Identifikasi Gap Penelitian	23
Table 4.1 Deskripsi Dataset	30
Table 4.2 Hasil Simulasi Dampak Promo Terhadap Revenue.....	50
Table 4.3 Karakteristik dan Label Segmentasi Pelanggan Hasil K-Means	55
Table 4.4 Evaluasi Performa Model Clustering.....	56
Table 4.5 Perbandingan Waktu Komputasi K-Means dan Hierarchical	57
Table 4.6 Segmentasi Pelanggan Hasil RFM dan Strategi Pemasarannya	59



DAFTAR GAMBAR

Gambar 3.1 Tahapan Penelitian	26
Gambar 4.1 Proses Membaca Dataset Transaksi Pelanggan	33
Gambar 4.2 Struktur dan Tipe Data Kolom pada Dataset Awal	33
Gambar 4.3 Proses Pembersihan Data	34
Gambar 4.4 Proses Pemeriksaan Dataset Setelah Pembersihan Data	35
Gambar 4.5 Proses Perhitungan Nilai Recency, Frequency, Monetary	36
Gambar 4.6 Proses Penyimpanan Dataset RFM	36
Gambar 4.7 Proses pembuatan visualisasi heatmap korelasi antar nilai	37
Gambar 4.8 Heatmap korelasi antar nilai RFM	37
Gambar 4.9 Proses Normalisasi Dataset RFM	38
Gambar 4.10 Metode Elbow untuk Menentukan Jumlah Cluster Optimal	39
Gambar 4.11 Grafik Metode Elbow untuk Menentukan Jumlah Cluster Optimal	40
Gambar 4.12 Silhouette Score untuk Menentukan Jumlah Cluster Optimal	41
Gambar 4.13 Grafik Silhouette Score terhadap Jumlah Cluster	42
Gambar 4.14 DBI untuk Menentukan Jumlah Cluster Optimal	42
Gambar 4.15 Grafik DBI terhadap Jumlah Cluster	43
Gambar 4.16 Proses Clustering Menggunakan K-Means	44
Gambar 4.17 Proses Reduksi Dimensi dan Penambahan Hasil PCA	45
Gambar 4.18 Proses Visualisasi Cluster Pelanggan dengan PCA	45
Gambar 4.19 Visualisasi Hasil Cluster Pelanggan dengan PCA	46
Gambar 4.20 Proses Clustering Menggunakan Hierarchical Clustering	47
Gambar 4.21 Proses Visualisasi Dendrogram Hasil Clustering	48
Gambar 4.22 Proses Penyimpanan Hasil Clustering ke File CSV	49
Gambar 4.23 Simulasi Dampak Promo Berdasarkan Segmentasi Cluster	49
Gambar 4.24 Proses Evaluasi Hasil K-Means Clustering	51
Gambar 4.25 Proses Evaluasi Hasil Hierarchical Clustering	51
Gambar 4.26 Visualisasi Hasil Segmentasi K-Means	53
Gambar 4.27 Hasil Karakteristik Tiap Cluster dari K-Means	53
Gambar 4.28 Visualisasi Hasil Segmentasi Hierarchical Clustering	55

DAFTAR LAMPIRAN

Lampiran 1 Kartu Asistensi	71
Lampiran 2 Curriculum Vitae	72
Lampiran 3 Surat Pernyataan HAKI	73
Lampiran 4 Sertifikat BNSP	75
Lampiran 5 Surat Ijin Riset Perusahaan.....	77
Lampiran 6 Form Revisi Dosen Penguji.....	78
Lampiran 7 Hasil Cek Turnitin	80
Lampiran 8 Lembar Persetujuan Sidang	81

